

DOCUMENT RESUME

ED 288 524

IR 052 139

AUTHOR Blackman, Bernard I.; Clevenger, Theodore, Jr.
TITLE Database Management Systems for Qualitative Research.
PUB DATE May 87
NOTE 22p.; Paper presented at the Annual Meeting of the International Communication Association (Montreal, Quebec, Canada, May 1987).
PUB TYPE Information Analyses (070) -- Reports - Descriptive (141) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Database Management Systems; Information Retrieval; Literature Reviews; Microcomputers; *Qualitative Research; Research Methodology; *Research Tools

ABSTRACT

This discussion of the development of database management systems specifically created to support qualitative research identifies two main problems in qualitative approaches, i.e., the issues of data collection and analysis, and the cumbersome nature of traditional qualitative approaches. These problems are reviewed with a comparison of the methods of manual data management and computerized data management approaches. The paper concludes with a discussion of several different uses of computer assistance in managing qualitative data and the identification of particular software programs for use in qualitative research. A 40-item reference list and a description of an integrated, menu-driven microcomputer system for managing text-based qualitative data--QUALPRO 3.1--are included. (RP)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

* This document has been reproduced as
received from the person or organization
originating it.

□ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

**DATABASE MANAGEMENT SYSTEMS
FOR
QUALITATIVE RESEARCH**

presented to

Human Communication Technology Special Interest Group

International Communication Association

Montreal, Canada

May, 1987

By

BERNARD I. BLACKMAN

and

THEODORE CLEVINGER, JR.

**COLLEGE OF COMMUNICATION
THE FLORIDA STATE UNIVERSITY
TALLAHASSEE, FLORIDA 32306**

(904) 644-5034

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY
Bernard J. Blackman

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

ED288524

IR052139

DATABASE MANAGEMENT SYSTEMS FOR QUALITATIVE RESEARCH

Abstract

As qualitative methodologies grow in popularity and respectability, there is an increased concern for reducing subjectivity and costs while improving consistency, dependability, and efficiency. The traditional manual or by-hand methods of managing data are idiosyncratic, anachronistic, and generally closed to public scrutiny. Moreover, manual data base management systems cannot confidently handle large volumes of textual data.

Recently, computers have been used to replace the mechanical functions of data management. Structured data base management systems (DBMS) are powerful tools for storing, indexing and retrieving scientific and business information, but they require a priori unitizing and data categories, thus lack the flexibility required for qualitative research. More recently, text based management systems (TBMS) have overcome some of these limitations, but depend upon whole-record retrieval models, thus fail to provide for the ad hoc approach to unitizing that characterizes much qualitative research.

Recognizing the limitations of commercial software designed for scientific or business applications, researchers have developed database management systems specifically to support qualitative research. These data base management systems facilitate the organization, storage, coding and retrieval of textual data. With varying levels of efficiency and effectiveness, these programs facilitate the processes of text coding and/or keyword/phrase searches. By clearly separating the mechanical functions required by qualitative methods from the analytical phase, such use of computers in textual data management can not only speed up the research process but also reduce the subjectivity and enhance the quality of qualitative studies.

Though still far short of the level of expert systems (let alone artificial intelligence) computer assisted systems for the management of qualitative data have grown progressively more useful, more accessible and easier to learn and use. The most recent addition, QUALPRO 3.0, is an integrated, menu-driven system for managing text-based qualitative data. It uses the microcomputer to facilitate, not replace, the researcher's own coding, indexing and retrieving processes for qualitative data.

DATABASE MANAGEMENT SYSTEMS FOR QUALITATIVE RESEARCH

INTRODUCTION

Empirical research based on nonnumeric or qualitative data is becoming increasingly popular in the study of human communication. (Kinzer, 1982) Long an established method in sociology, indeed the dominant method in anthropology, qualitative method has recently been embraced by education (Fetterman, 1984; Goetz & LeCompte, 1984; Wilson, 1977), management, organization, and communication research. Special issues in Administrative Science Quarterly (1979, 1983), Journal of Management Studies (1982, 1983), and Western Journal of Speech Communication (1983), as well as recent books (Morgan, 1983; Putnam & Pacanowsky, 1983; Van Maanen, 1982) on symbolic or interpretive approaches indicate a renewed interest in qualitative methods.

In spite of the profusion of qualitative approaches, critics and advocates alike voice concern over the anachronistic and idiosyncratic nature of the research as practiced today. Further development of qualitative method may depend on finding ways to improve the adequacy of data collection, the consistency and thoroughness of analysis, and the dependability of results.

One part of the problem is that reports of qualitative research rarely address issues of data collection and analysis (Lofland, 1974). Without attention to such questions the evaluation of qualitative studies rests on critical assessment of the "authenticity" or "reasonableness" of the author's description of the results. Too often this seems more like faith than like critical and objective evaluation.

Recently scholars have begun to demystify these processes through explicit articulation and formalization of qualitative data management procedures (Bogdan & Biklen, 1982; Bogdan & Taylor, 1975; Glaser, 1978; Goetz & LeCompte, 1984; Miles & Huberman, 1984; Patton, 1984; Spradley, 1979, 1980). The object of this explication is summed up by Miles & Huberman (1984), who state that "the field of qualitative research badly needs explicit, systematic methods for drawing conclusions, and for testing them carefully" (p.16).

A second aspect of the problem is the awkward and cumbersome nature of traditional qualitative methods. Generally speaking, data have been collected, stored, retrieved, transformed and interpreted "by hand". Much of the work is essentially clerical. This mechanical form of data management is inherently tedious, slow and unreliable when done by human hand, but can often be done better and more easily by a computer. Although the use of computers in quantitative field research is well developed (Kirk, 1981), the idea of applying computers to qualitative data is still in its infancy.

The issue is straightforward - improve objectivity, open the procedures to more public scrutiny, introduce more rigorous standards, and install computers in qualitative research methodology. Given the need for computers to assume some of the more mundane and repetitive tasks in qualitative research, we must begin to identify the tasks involved. It is equally important that while the method is being transformed, the unique contribution that the subjective researcher brings to the process of exploration and analysis not only be preserved, but be carefully articulated and strengthened. This is most critical as technological innovations destroy the narrowly confined traditional approaches to data accessibility and management.

At least four specific changes in data management and communication will forever transform how research is approached, conceptualized and conducted. First, as microcomputers become more commonplace data will become more oriented to disk (and other electronic media) than paper. Documents available on disk are readily accessible to broader and more diverse audiences and amenable to reorganization and compilation.

Second, for those materials that are not already on disk, the transfer from manuscript to disk is becoming faster, easier and cheaper. Optical scanners can now read and convert a page of text directly into a disk file in a brief time.

Third, optical compact disks, like CD-ROM and CD-WORM, increase the storage capacity of a floppy disk medium hundreds of times. That means, for example, that most if not all of ancient Greek literature can be stored on one or two disks the size of a 5 1/4 inch floppy disk and used on a microcomputer at home.

Finally, intersecting the other three innovations is the linkage of microcomputers into networks that increase the opportunity for information to be readily shared and reorganized. These four technological innovations, disk based text, optical scanners, communication networks, and compact laser disks, have important implications for how qualitative research will be conducted in the future. This paper explores the convergence of microcomputing, enhanced data storage and management, qualitative methods and communication research by asking how computers can be used to manage data in qualitative research.

This paper will focus on extant technology, not theoretical potentials. First, to clarify the problems that computer systems must address, we briefly review manual data management traditions and procedures. Second, we discuss the recent trend in computer-assisted data base management programs appropriate for textual data. Finally, we describe QUALPRO, a data base management program and productivity tool designed specifically for qualitative research.

MANUAL DATA MANAGEMENT

Because of the researcher's active involvement in every phase of the process, qualitative research is often perceived as subjective, unsystematic, and of limited generalizability. Dennis, Goldhaber, and Yates (1979) suggest that lack of systematization, and therefore lack of replicability, rob the method of its "potentially useful and conceptually rich form of analysis" and reduce it principally to an "anecdotal source of Gestalt impressions" (p.260).

The undeniable subjectivity of qualitative research rests in part on data management systems which are potentially unsystematic, highly personal, situation-specific, and time consuming. Using such a system with very large data bases greatly increases the likelihood of error. In fact, Agar (1983) has called the management of the immense volume of "non-numeric records - field notes, transcribed informal interviews, documents, genealogies and so on," the classic or core problem in qualitative studies (p.19).

To understand why data management is crucial to the quality of the product, consider how qualitative research is typically done. First, data must be collected "in the field" and recorded in some way, usually in written form. For every hour in the field, the researcher may spend up to four or five hours writing up notes and observations (Simmons, 1969; Scudder & Colson, 1979). Because the research calls for the most "complete, accurate, detailed field notes" possible, this stage of the process often produces enormous volumes of data. This in turn requires a tremendous investment of the researcher's time in data management and mandates "self-discipline, if not compulsion" to do a consistent, systematic and thorough job (Bogdan & Taylor, 1975, p.60).

As the data are gathered, they must be organized and then constantly reorganized as new data are added. With each increment the organization task grows in size and complexity. It is not uncommon for the major time investment to be in the organization of data. Roth (1974) and Werner (1982) claim that organizing data may take months or years, whereas writing the report may take only a few weeks. This may draw the study out over a long period of time, which greatly reduces the usefulness of qualitative methods as a companion to quantitative designs in team or multi-method projects (Miles & Huberman, 1984).

Even as the data are being gathered, stored and organized, the next phase of the research begins; the data must be analyzed and interpreted. This process will test the limits of the researcher's capacity for handling large volumes of data. Indexing, coding or classification systems are designed and applied to the data base. Depending on how well they work, they are supplemented, modified or replaced. At each step the data must be recoded and analyzed anew. Pelto (1970) says that "beyond a certain point the tasks of storing and manipulating these data become so enormous that we tend to lose track of our materials, or our inventorying and managing of the data become overly cumbersome and time consuming" (p.356).

These observations make it clear that managing the data base is a major bottleneck in qualitative research. The researcher's attempt to be systematic and thorough in data collection will most likely lead to larger collections of data that

will be cumbersome, and require more time to manage. As the database grows arithmetically, the difficulty of managing it grows geometrically, with potentially devastating effect.

Because of the geometric ratio between the size of a database and the difficulty of managing it, the relationship between the amount of data and the quality of the research may be curvilinear. Under traditional methods of data management, there comes a point after which the probability increases with the amount of material collected that consistency and thoroughness toward the data will slacken, with a concomitant loss of quality.

Recognizing the centrality of the data management problem, researchers have sought solutions. One approach has been to put limits on the scope, depth or length of data collection. Scudder and Colson (1979) recommend that to avoid the dangers associated with too much data the researcher should intentionally limit the study a priori so as to be "sure that one does not get swamped in a mass of information which cannot be stored in the mind and thought about" (p.249). Also, Wiseman (1974) described fragmentation of data analysis into smaller, more manageable portions as a means to manage a large volume of data. However practical these solutions, they minimize the quality of a study because they impose restrictions based on labor costs or data management capacity rather than on the degree to which the original research questions have been satisfactorily answered.

Researchers have noted that working with a large data base encourages more human errors in every aspect of the work: writing up notes, coding and analyzing texts, and writing reports (Scudder and Colson 1979). The effect of long hours of dull and boring work may be errors, omissions, overlooked themes and areas of investigation to pursue, short-cut methods, and superficial or inconsistent cross-checking of facts. As Roth (1974) points out, the addition of time saving devices like dictation-transcription equipment or the use of a typist, while clearly reducing the workload, may still compromise quality of results because they remove the researcher from direct and immediate contact with his/her notes. In a worst-case scenario, failure to solve data management problems may leave large amounts of data unanalyzed and uncoded (Goetz and LeCompte, 1981).

Qualitative method is not inherently unsystematic or inefficient. However, "by-hand" methods of data management in qualitative methods, particularly in the repetitive, mechanical aspects of manipulating large volumes of data, lead to serious questions about the consistency and value of the data and the conclusions of the study.

In an effort to systematize qualitative data management, recent publications describe various data management and analytic schemes to achieve, as Becker (1958) said, "more a 'scientific' and less an 'artistic' kind of endeavor" (p.660). Miles and Huberman's (1983) book on qualitative data analysis is but one attempt to increase awareness and scrutiny of the analytic process by which textual data are transformed into concepts and conclusions.

Any method of storing and retrieving data in a systematic way can be called a data base management system (DBMS). Such systems have several common objectives in qualitative methods. They economize time working with textual data

and bring order to the physical manipulation the texts. Specifically, a DBMS should:

1. Store data easily and efficiently,
2. Retrieve the complete, unmodified text at any time,
3. Allow for flexible coding within the text of units of any size, which are easily sorted into separate categories,
4. Facilitate the retrieval of code names, code frequencies, and textual units by code or category.

Traditional DBMSs in qualitative methods have been manual systems developed by the individual craftsman in response to his particular needs in a study, i.e., idiosyncratic to the researcher and situation. However, a few general types can be distinguished. Bogdan and Biklen (1982) describe four ways data can be managed in a mechanical, by-hand fashion.

The first method, "eyeballing," which they do not recommend, relies simply on scanning the text and writing from memory. With this system the core storage is the original documents with retrieval dependent on the limits of the researcher's memory.

In the second DBMS, a file folder system is used to facilitate mechanical storage and expand the researcher's memory capacity. In this system, called the "cut-up-and-put-in-folders" approach, the researcher keeps the original field notes complete while cutting up copies of the field notes into coded units of data. The cut up units are then sorted into folders representing different categories (Bogdan & Taylor, 1982). Any category is retrieved by going to the folder holding the units of data with that category name.

Card filing represents the third type of DBMS. In this system, identifying labels including the name of the code, and the page and line location in the text are put on cards. Units of data classified by a code are located by searching the cards and then locating the appropriate unit in the text. An alternative method, suggested by Miles and Huberman (1984), puts coded chunks of data directly on each card.

The fourth system is a commercially available punched-card system. With this method, each unit of data is recorded on a card with pre-punched holes around the margin. The researcher notches appropriate holes corresponding to codes in the top of the cards. Retrieval of a code category is handled by assembling the cards into a deck, pushing a needle through the holes of the desired category and extracting from the deck all cards with the appropriate code.

While each of these last two types of DBMS is an improvement over the "eyeballing" approach, they are still labor-intensive, slow, and generally inefficient. With the file card approach, for example, it is difficult to skim the notes, retrieval remains laborious and may require resorting pages after each retrieval, and the method does not facilitate handling cross-coded units of data (Bogdan &

Biklen, 1982). The cut and file approach usually requires multiple copies of the text, increases the handling of many pieces or fragments of paper, and usually requires a large space to sort the codes.

Though each of the DBMS described by Bogdan & Biklen (1982) suffers from major limitations, each brings some structure and systematization to data handling and may be used effectively with small data bases. (For an explicit and detailed description of the procedures in a study using a manual DBMS see Wiseman, 1974). But with moderate or large data bases, serious questions arise about the reliability of any of these systems due to the central involvement of the researcher in the data management process.

COMPUTERS AND QUALITATIVE RESEARCH

For large data bases, much might be gained by transferring the burden of data management from the hands of the researcher to the computer. The problems that plague research on large qualitative data bases would be overcome, and research could be made more systematic by formalizing the unitizing, storage, retrieval, classification and sorting of data. Two problems have previously stood in the way of computer use by qualitative researchers; one problem is perceptual, the other logistic.

The perceptual problem has to do with doubts about the computer's ability to contribute to an essentially interpretive process. Understandably some researchers feel that mixing technology and interpretation will mitigate creativity or dehumanize the analysis. However, this fear is unwarranted because the mechanical and interpretive aspects of data management are distinct and separable. This distinction holds even though data collection, management and analysis occur simultaneously throughout a qualitative study. Dass (1983), for example, says that

interpretive work requires the data manipulation of the mechanical operations yet, at the same time, the interpretive will dictate the type and, in many cases, the content of mechanical operations. However, despite the fact that each phase informs the other, their respective operations are separate and distinct. (p.337)

This distinction between the mechanical and the interpretive activities is most readily appreciated by comparing various computer applications in content analysis. Generally, computers have been applied in two ways: computerized content analysis and computer-aided analysis. Computerized content analysis attempts to automate the processing of interpretive data, while computer-aided analysis yields data presumed useful to interpretation by a human analyst. According to Krippendorff (1980), in computerized content analysis the

computer is programmed to mimic, model, replicate, or represent some aspect of the social content of the data it processes, that is, where the process of drawing inferences from original text (or data generally) is entirely or largely computerized. (p.123)

Basically this is accomplished either by word frequency counts or programs that compare words in the object text to a dictionary, and then code according to a priori instructions. ANACONDA (Bierschenk, 1977) is one program used in the analysis of textual data. More recently, Patton (1984) used a computer to help analyze a large data base. He first developed a comprehensive category system and coded every paragraph of every page on the collected data. Computer cards were punched for every idea in every paragraph, with information that identified the unit of code. Once keypunching was completed and debugged, units bearing any predetermined code could be retrieved from the entire data base and printed.

In these examples of computerized content analysis, the analysis is not truly automatic. The computer's data processing is limited to a priori categories defined by the programmer. Programs may aid categorization and quantification through superhuman speed, but they cannot yet replace human symbolic processing (Agar, 1983). Thus even this form of "automated" content analysis reflects the distinction between clerical and interpretive functions in analyzing symbolic data bases.

The distinction is even sharper in the case of computational aids, procedures employed before or concurrent with data conceptualization and analysis. Krippendorff (1980) says they give the researcher an "overview of the variety, kind, and distribution of data" (p.121). With computational aids computers do not analyze content, but rather assist the researcher in preparing for analysis. Krippendorff (1980) says computational aids include alphabetical lists, word frequencies, lists of selected words in their linguistic context, and information retrieval programs designed to "search and locate" bits of information.

Information retrieval programs (the name given early types of DBMS) can facilitate interpretation by reducing large volumes of data to manageable portions. In information retrieval programs the researcher directs the program to select the type of information he/she wants and then interprets the selected information. Therefore, the researcher applies human analytic skills at three levels: developing codes on the raw data, directing the program through the data base, and interpreting the retrieved information.

As these examples illustrate, the separation of data management into mechanical and interpretive processing, with computers doing the mechanical and humans the interpretive tasks, maintains human involvement in the research process while reducing many of the serious threats to the quality of the results by minimizing the errors usually associated with the subjectivity of the process. In fact, with computer assistance, the researcher not only processes data more systematically and thoroughly, but has time for more complete and exhaustive data collection. The computer can remove data management limitations on the scope,

depth or length of a study and enable the researcher to handle more data more easily, quickly, and reliably.

The second problem standing in the way of computer use by qualitative researchers has been logistic. Until quite recently no computer program has been available for general use in qualitative research. The computer methods of content analysis described above are implemented by mainframe software requiring a programmer's approach. Once mastered, they rarely fit either the data or the analytic framework of the qualitative researcher.

Database management systems available for the microcomputer are beginning to be used to assist qualitative research. While some programs may be valuable aids to qualitative researchers, most are insufficient for their specialized needs. We can better understand the problem of using the computer in qualitative research if we first explore why even the best of the traditional microcomputer DBMS may fall short of the specialized needs of the qualitative researcher.

Richard Byrne (1985) captures the contemporary concept of DBMS in the following:

A data base is a collection of related data, names, numbers or images, stored in such a way that the data can be organized, reorganized, selected and presented for some specific purpose. (p. 25)

Byrne's examples of data bases are telephone directories, mail-order catalogs, dictionaries or a checkbook. These fit the definition because they are organized alphabetically, numerically or chronologically, to help find information quickly. A typical counter-example to the traditional data base is a newspaper or magazine. It has lots of information, which is organized under broad headings, but the data are not indexed for rapid retrieval and summary. The qualitative researcher's "data base" is much more like a newspaper than it is like a telephone book. In fact, it is more like a mixed collection of newspapers, magazines and old grocery lists.

What distinguishes traditional data bases from the data gathered by qualitative researchers is the level of a priori organization. In gathering data for a phone book the compilers know in advance just what information will be needed: last name, first name, street address, and telephone number. Everything else being superfluous, it is easy to organize the information into records (one for each patron) and fields (one for each type of data, such as last name), and to arrange these in an orderly arrangement conducive to rapid sorting, reorganizing and retrieval of any individual datum or defined set of data.

The problem is not that data bases are restricted to short records. One popular program for microcomputers, PARADOX, can handle fields up to 4000 bytes, roughly two single-spaced typewritten pages in length. It accommodates 225 such fields per record, and up to 65,000 such records per file (Coward, 1985). This adds up to more raw capacity than the qualitative researcher is likely to need, but it is power that cannot be harnessed to the needs of qualitative research.

Nor is the problem that data bases provide limited ways to search the records. Some DBMSs provide virtually limitless search and sort capabilities from

category codes through key-word searches to complex Boolean procedures. For example, any of a dozen popular DBMSs for the microcomputer could take a command of the form: "Print out the full records for all of our salesmen who have been with the company more than two years, worked in only one product line, sold exclusively in markets of 250,000 to 500,000 population and had average annual sales exceeding \$2 million." The program could scan a data base of several hundred salesmen and be ready to print out the answer in a surprisingly brief time, determined by the hardware and operating system of the computer. Using keywords and Boolean instructions, a DBMS can duplicate many of the forms of computerized and computer-aided content analysis described above, given data input in the appropriate format.

Indeed, a full-featured business or scientific DBMS can answer almost any question one can think of, provided the question is phrased in terms of predefined fields; that is, fields that were defined before the data were entered. Therein lies the problem for the qualitative researcher, for it is very often the case in qualitative research that the categories and classifications are not known before the researcher goes into the field. Indeed, one of the principal objects of qualitative research is to discover the structure of the phenomenon under investigation, and the structure of the phenomenon determines the structure of the data base and the coding categories.

The data structure for most qualitative research is not an a priori framework into which the data are filed; rather it emerges from the data as they are collected. It usually grows and transmutes in unpredictable directions during the research. Therefore, a traditional DBMS is applicable to the data of a qualitative study, if at all, only after the creative and analytical phases of the study are finished or if the data are entered in highly structured pre-defined fields.

A second limitation of ordinary DBMS has to do with the problem of unitizing. The typical DBMS requires that data be entered into fields as intact units. For example, an annotated bibliography in a data base might have fields for the author's name, title, journal and date. Additional fields might place the entry in one or more topical categories, plus a field containing an abstract or annotation. Any entry (or set of entries) might be retrieved or the whole data base sorted very quickly by author, journal, topic, date, keywords in the abstract, or any combination thereof.

The price for this data management power is standardization. Each field must contain homogeneous information units of similar type and complexity. This is not a serious problem in most business applications or quantitative research projects, where a priori standardization of units is essential. However, it is a critical problem for qualitative research, where an important goal may be to answer the question: "With respect to this problem, what constitutes a meaningful unit of data?" The answer may lead to multiple levels of analysis with new layers of hierarchical structure emerging as the project evolves. Typical data base management systems cannot accommodate this type of complexity or serendipity.

Though the structured DBMSs developed for business and scientific data management seem to promise help to the qualitative researcher, in fact they have limited application. Because they require a priori data unitization and category

definitions, they cannot accommodate the emergent character of the qualitative research process.

In response to the limitations imposed by structured DBMSs, a number of software vendors have developed Text-Based Management Systems (TBMS) for microcomputers. These allow the user to define a database with fields of variable length that can grow dynamically. These programs generally require that data be stored in the form of whole "records" which may not be subdivided into sections or parts. Each record may then be coded for a large number of variables. Many TBMSs provide for multiple coding within a given variable.

The database is searched or indexed by examining the codes or keywords attached to each record and/or by a whole-text search (scrutinizing the document word-by-word for desired words or phrases). Such programs usually excel at boolean searches, quickly retrieving or identifying records that display complex combinations of codes and content.

Excellent TBMS programs are available in a wide range of prices. At the low end, SquareNote from Union Squareware is available from discount distributors for under \$50. Each record has a maximum length of 500 lines, consisting of three fields: keywords, header and body. The keyword list on a record can be expanded at any time. Records may be retrieved using the keywords or by full-text search. The program does not provide for true boolean searches, but a simple boolean effect can be achieved through ingenious programming.

Farther up the scale of both utility and cost, AskSAM from Seaside Software offers ease of use, power and flexibility at a price of \$200. Each record has a 20-line limit, but records may be linked together into "documents" of any length. Creating a database in AskSAM requires only that the record be typed in: no field names, no field typing or sizing is required. Individual records may have specified fields, but they need not, and records with different numbers and types of fields may co-exist in the same database. Searches can be simple or as complicated as necessary. The program has a highly structured command language that includes boolean and comparison operators.

Puglia (1986) assessed the state of the art in TBMS and reviewed in detail ten programs, including the two above, covering a price range from \$50 to \$1,495 retail. A common limitation is that these programs all retrieve or index only whole "records" or, at most, the contents of defined fields. While these operations are sometimes useful in qualitative research, they do not lie at the heart of qualitative method. Thus, TBMSs probably hold a potential for greater contribution to qualitative research than structured DBMSs, but they are not likely to help greatly in the areas where help is need most, coding and organizing textual units of variable length embedded in numerous documents.

The type of DBMS that would best serve the needs of the qualitative researcher would allow for entry of whole-text documents of any type, length or format. It would provide for ad hoc category definitions, continuously modifiable during the life of the project. And it would allow for flexible unitization; that is, it would allow the researcher to define and redefine units of any length, in any relationship to other units, at any stage of the project.

In principle, of course, the computer has the potential to offer this flexibility and much more. One could imagine an integrated expert system capable not only of doing word searches and frequency counts, but also a variety of symbolic processing functions such as parsing, thematic content analysis and pattern recognition. In principle, such a system might ultimately be brought to the level of artificial intelligence, where it could perform certain interpretive tasks as well.

Regrettably, such a system is not available now nor is it likely for some time. Although sophisticated systems are available for parsing and pattern recognition, decision rules for applying these in qualitative research are wanting. This is hardly surprising, since the problem of decision rules remains one of the most confused and hotly contested issues in qualitative analysis.

Happily, even in the absence of well-accepted decision rules for converting text into coherent data, qualitative researchers are able to use computers in a variety of ways to ease their workload and reduce the likelihood of error. We turn next to several computer DBMSs that have approached this ideal to varying degrees.

COMPUTERIZED DATABASE MANAGEMENT PROGRAMS FOR QUALITATIVE RESEARCH

Several different uses of computer assistance in managing qualitative data have recently appeared. Some, like Agar (1983), recommend using microcomputers as field tools in managing textual records, and others rely on mainframe computers for expanded memory capacity (Burnett, 1972; Blackman & Provencher, 1984; Drass, 1980). Now that commercial text databases for microcomputers are more readily available they are slowly being adopted to replace many of the manual operations in qualitative research. Moreover, programs specifically designed for qualitative research are being written.

Described below are several computer programs for managing qualitative data. Each of these DBMSs attempts to provide a novel solution to the task of managing the large volume of data while reducing the costs associated with the manual data handling. Although each of the programs appears to have similar objectives, they differ in how well and how significantly they relieve the researcher of the mechanical data management tasks, and how easily they are learned and used.

Bernard and Evans (1983) describe a system for field notes that combines commercially available text editors and DBMS. The text editor facilitates field note entry and prints hard copies of the field notes. The DBMS stores codes with the page numbers of their locations in the field notes. When a code is retrieved, the location of the page where the codes appear are printed. While this system does facilitate data management, it is limited. In brief, while data entry and locating codes are clearly easier, the retrieval, compilation and print out of codes is still manual.

Werner (1982) has created a flexible indexing and information retrieval system that combines several available software packages (i.e. electric blackboard and windowing) with his own program. Key words are used to extract small portions of text.

Sproull and Sproull's (1982) TEXTAN program combines a text editor with a specifically designed program that analyzes text data. Natural language codes are embedded directly into files in an attempt to preserve the context of codes.

These three systems illustrate some of the novel attempts to develop a flexible DBMS for qualitative data. However, while these programs are innovative solutions, they either appear difficult to use and learn, are too specialized to a particular type of research purposes, or use codes embedded into the text itself.

The next three systems, BEDRES (Burnett, 1972), LISPQUAL (1980), and QUAL*PRO, 1.0 (Blackman & Provencher, 1984) were specifically designed to process large textual data bases in ethnographic studies. Each of these systems appears to demonstrate an increase in storage capacity, flexibility of structure, and efficiency in operation.

BEDRES (Burnett, 1972) is a package of programs employed serially, very much like the old SPSS punch card system, to perform the mechanical functions of data management.

LISPQUAL (Drass, 1980) is a fully interactive command-driven program in which each command performs a different mechanical function on the data. By combining different commands, the researcher can search and retrieve and code data while maintaining the integrity of the original field notes.

QUAL*PRO, 1.0 is also an interactive system which is designed to facilitate the text entry, coding and retrieval of coded text.

These three programs share a common conceptual design and purpose, but they differ somewhat in how the design is enacted. The programs can store large textual data bases, search and retrieve data efficiently, attach or remove codes without altering the original field notes, and obtain code frequency data. All three programs use natural language codes and can retrieve codes from the data base along with identifying and contextual information.

While BEDRES, LISPQUAL, and QUAL*PRO, 1.0 effectively transfer many of the repetitive mechanical functions associated with qualitative data management to computers, they remain awkward to operate. BEDRES uses a primitive command card operation that seems lengthy and error-prone. LISPQUAL requires the user to learn and use 52 functions in combinations in order to operate the system. And QUAL*PRO, 1.0 requires a special TUTOR program to learn its operations.

In response to the need for a more user-friendly system, QUALPRO, 2.0 and 3.1 were developed. QUALPRO, 2.0 is a fully interactive, menu-driven program written in BASIC that performs all the operations of the DBMS's discussed above, but in a simple, user-friendly manner. No programming and little computer

knowledge beyond signing on to the computer and text editing are required. Although QUALPRO 2.0 is a valuable research tool, it operates only on a mainframe computer.

QUALPRO 3.1 is a menu driven and user-friendly productivity tool and text database manager designed for IBM PC (and compatible) microcomputers. When QUALPRO is used as a text database manager, the researcher can structure and label text documents; once text documents are entered and organized, the database can be searched and documents easily retrieved. When QUALPRO 3.1 is used as a productivity tool codes (or keys) can be electronically attached to sections of text within documents. The coded sections can be located, copied, labeled, and directed to the screen, a printer or textfile. Each extracted section of text is printed with information identifying the document and its location within the document.

The history of text database managers shows a pattern of evolution toward broader utility and ease of use. The latest of these programs, QUALPRO 3.1, is also the most advanced. It parallels closely the manual approach to coding and retrieving text but greatly accelerates the process and enhances its reliability. A fuller description of QUALPRO 3.1 is included in Appendix A at the end of this paper.

Each of the systems described above has certain advantages, and collectively they represent progress toward greater utility as well as a more "intuitive" look and feel. Even so, none approaches the ideal of the intelligent system described in the previous section. As text database managers for qualitative research evolve toward expert (and ultimately intelligent) systems, the next steps should be (1) incorporating both keyword/phrase searches and coded-unit searches in the same program and (2) providing for boolean searches which accept as parameters both keywords/phrases and coded units either separately or interactively.

Qualitative researchers should keep a sharp eye on the commercial market of business-oriented text-based data management systems, for these are growing more powerful and flexible with each generation. However, because of the unique demands of qualitative research, the software best tailored to the needs of ethnographers, critics and qualitative content analysts probably will continue to emanate from the research community itself.

In summary, Qualitative research is increasingly popular in the study of human communication, and holds great promise for advancing our knowledge about the communication process. However, the method stands in need of improvement. Many of the limitations that plague the process of qualitative research are precisely those limitations of human frailty that computers redress so competently. Furthermore, four changes in computer technology -- disk-based text, optical scanners, communication networks and compact laser disks -- as well as the proliferation of powerful microcomputers are permanently altering the way we deal with qualitative data.

We need software able to match this technology. Each of the programs described in this paper has advanced the application of this powerful technology to data base management systems for qualitative research. They represent a growing trend toward using computers to assist the researcher by simplifying the mechanical

functions of processing data. Computer-assisted programs can bring more organization, consistency and efficiency to the management of data than traditional manual methods. With computer-assisted data management, qualitative studies can be pursued more broadly and in greater depth, and at the same time be cheaper, more thorough and less prone to clerical error.

The use of computerized data base management systems will help to bring more public scrutiny to what has traditionally been a rather arcane process. By expanding limited human capacities for the mechanical aspects of data management and freeing and facilitating the human capacity for analysis, the nature of the process may be explicated and the quality of studies improved. The consequence of merging computers into qualitative research may be more and better communication among communication researchers.

REFERENCES

- Agar, M. (1983). Microcomputers as field tools. Computers and the Humanities, 17, 19-26.
- Becker, H.S. (1958). Problems of inference and proof in participant observation. American Sociological Review, 23(6), 652-660.
- Bernard, H.R. & Evans, M.J. (1983). New microcomputer techniques for anthropologists. Human Organization, 42(2), 182-185.
- Bierschenk, B. (1977). A computer-based content analysis of interview texts: Numeric description and multivariate analysis. Didakometry, 53.
- Blackman, B.I. & Provencher, D. (1984). Computer assistance for qualitative methods in intercultural communication research. Paper. International Communication Association. San Francisco.
- Bogdan, R.C. & Taylor, S.J. (1975). Introduction to Qualitative Research Methods. New York: John Wiley.
- Bogdan, R.C. & Biklen, S.K. (1982). Qualitative Research for Education. Boston: Allyn & Bacon.
- Burnett, J.H. (Ed.). (1972). Computer assisted processing of ethnographic data. (Report). University of Illinois: Bureau of Educational Research.
- Byrne, R. R. (1985). The well managed database. Personal Computing, October, 9(10), 25+.
- Cowart, R. (1985). Paradox from Ansa software: High-quality database program is fast, easy. PCWeek, November 12, 1985, 23-128.
- Das, T.H. (1983). Qualitative research in organizational behavior. Journal of Management Studies, 20(3), 301-314.
- Dennis, H.S., Goldhaber, G.M. & Yates, M.P. (1978). Organizational communication theory and research: An overview. In B.D. Ruben (Ed.), Communication Yearbook. New Jersey: Transaction Books.
- Drass, K.A. (1980). The Analysis of qualitative data: A computer program. Urban Life and Culture, 9(3), 332-353.
- Fetterman, D.M. (Ed.). (1984). Ethnography in educational research. Beverly Hills, CA: Sage.
- Glaser, B.G. (1978). Theoretical sensitivity. Mill Valley, CA: Sociology Press.

- Goetz, J.P. & LeCompte, M.D. (1984). Ethnography and qualitative design in education research. Orlando, FL: Academic Press.
- Goetz, J.P. & LeCompte, M.D. (1981). Ethnographic research and the problem of data reduction. Anthropology and Education Quarterly, 12(1), 51-70.
- Kinzer, Harold J. (Ed.). (1982). Uses of Qualitative Research: A Symposium. The Communicator. Journal of the Northwest Communication Association, 12(Spring), 3-95.
- Kirk, R.C. (1981). Microcomputers in anthropological research. In D.R. Heise (Ed.), Microcomputers in social research (pp. 473-492). Beverly Hills, CA: Sage.
- Krippendorff, K. (1980). Content analysis. Beverly Hills, CA: Sage.
- Lofland, J. (Ed.). (1974). Analyzing qualitative data: First person accounts. Urban Life & Culture, 3(3), 307-309.
- McCall, G.J. & Simmons, J.L. (1969). Issues in participant observation. Reading, Mass: Addison-Wesley.
- Miles, M.B. & Huberman, A.M. (1984). Qualitative data analysis. Beverly Hills, CA: Sage.
- Morgan, G. (Ed.) (1983). Beyond method. Beverly Hills, CA: Sage.
- Pacanowsky, M.E. & Putnam, L.L. (Eds.). (1982). Interpretive approaches to the study of communication. Western Journal of Communication, 46(2).
- Patton, M.Q. (1980). Qualitative evaluation methods. Beverly Hills, CA: Sage.
- Pelto, P.J. (1970). Anthropological research. New York: Harper & Row.
- Puglia, V. TBMS: database power unleashed. PC Magazine, 5(20), 211+.
- Putnam, L.L. & Pacanowsky, M.E. (Eds.). (1983). Communication and organizations. Beverly Hills, CA: Sage.
- Roth, J.A. (1974). Turning adversity to account. Urban Life & Culture, 3(3), 347-359.
- Scudder, T. & Colson, E. (1979). Long term research in Gwembe Valley, Zambia. In G.M. Foster, T. Scudder, E. Colson, & R.V. Kemper. Long term field research. New York: Academic Press.
- Spradley, J.P. (1980). Participant observation. New York: Holt, Rinehart & Winston.
- Sproull, L.S. & Sproull, R.F. (1982). Managing and analyzing behavioral records: Explorations in nonnumeric data analysis. Human Organization, 41(4), 283-290.

- Starbuck, W.H. (Ed.). (1982). Organizations as ideological systems. Journal of Management Studies, 19(1).
- Turner, B.A. (1983). The use of grounded theory for the qualitative analysis of organizational behavior. Journal of Management Studies, 20(3), 333-347.
- Van Maanen, J. (Ed.) (1979). Qualitative methodology. Administrative Science Quarterly, 24(4).
- Werner, O. (1982). Microcomputers in cultural anthropology: APL programs for qualitative analysis. BYTE, 7(7), 250.
- Wilson, S. (1977). The use of ethnographic techniques in educational research. Review of Educational Research, 47(1), 245-265.
- Winograd, E.B. and Winograd, G.R. SEPARATE.BAS: A microcomputer program to assist in the textual analysis of qualitative data. The Western Journal of Speech Communication, 51(1), 52-67.
- Wiseman, J.P. (1974). The research web. Urban Life and Culture, 3(3), 317-328.

APPENDIX A

DESCRIPTION OF QUALPRO 3.1

WHAT IS QUALPRO?

QUALPRO is a menu driven and user-friendly productivity tool and text database manager. As a Database documents can be imported or created in QUALPRO and then easily located and retrieved. When you use QUALPRO as a Productivity Tool you can electronically attach codes (or keys) to sections of text within documents. The coded sections can be located, copied, labeled, and directed to the screen, a printer or textfile. Each extracted section of text is printed with information identifying the document and its location within the document.

HOW DOES QUALPRO OPERATE?

QUALPRO runs under BASIC under DOS on IBM PC's, XT's and AT's and compatibles. QUALPRO uses less than 40K RAM but can manipulate files up to maximum disk size (360K on standard floppy disks or 1.2 MB on a high-intensity AT drive) by employing spare space on the system disk as a temporary file (buffer). On hard disk systems, file size is limited only by the available space on the hard disk.

WHO SHOULD USE THIS PROGRAM?

QUALPRO was written to facilitate the clerical and administrative tasks in qualitative research and therefore appeals to a broad range of researchers and writers. Because content analysis, participant observation, historical critical analysis and many other forms of field research depend on large amounts of text data, QUALPRO is an invaluable research assistant in these research approaches.

Anthropologists, psychologists, and sociologists who collect and manage large amounts of qualitative data will find QUALPRO an indispensable tool to assist them in the data management and interpretation process. Evaluation researchers in education and other fields of practice will find QUALPRO a valuable tool as well. In fact, anyone who uses large amounts of disk-based text can enhance productivity in writing and organizing writing through QUALPRO.

HOW DOES QUALPRO DIFFER FROM A TRADITIONAL DBMS?

Traditional database systems (DBMS) usually require prior and fixed-field designation before information can be retrieved. For example, name and address field types and sizes are determined before data entry, which then restricts the amount and type of data entry and the type of output.

QUALPRO was designed as a simple and flexible text database manager. As a Text Database Manager, QUALPRO organizes and retrieves whole documents without requiring field designation prior to data entry. Also, QUALPRO does not

set prior limits on the size of text documents. Therefore, retrieval of data is not limited to predetermined sort categories or keys.

Codes (or keys) can be electronically attached to a text document without altering the document. Any section of a document, down to a single line of text, can be located, extracted and reproduced. QUALPRO can electronically attach any number of different codes any number of times to any number of documents in a database. Codes and code frequencies can be compiled from a single document or from the entire database of documents.

Therefore, QUALPRO's unique advantage over the traditional databases is its flexibility in retrieving information from text in a manner that mirrors and supports the iterative process involved in the creative and interpretive processing of qualitative data. QUALPRO reduces the labor-intensive clerical aspects of managing data while permitting re-examination, exploration, and interpretation without burdening the user with clerical baggage.

HOW IS THE QUALPRO OUTPUT PRESENTED AND USED?

Information from a single document file, from multiple documents or from the entire database of documents can be directed to the screen, to the printer or to a textfile. Output can include HEADINGS, information identifying the document; FILES, the text document; CODES, sections of text extracted from the whole text document by key words; CODES SUMMARIES, information on code names; and CODE FREQUENCIES, information on the number of times a code appears.

For example, one code can be collected for the entire database. This produces displays (to the screen, textfile or printer) of copies of every occurrence of a code in the entire database. This is equivalent to a file folder of items cut from copies of documents and sorted into files, and is indispensable for interpreting the text in qualitative research and re-organizing the most important parts of documents.

Used in this manner, QUALPRO is very much like traditional library research where the researcher extracts quotes from different books and journals to be organized into a new pattern. In fact, combining QUALPRO with other microcomputer programs automates many of the manual parts of the research process. For example, the compiled output files from QUALPRO can be put into an outline processor such as THINK TANK or READY! and reorganized to represent some topical themes. When edited in a word-processor, such as Microsoft WORD (2.0 and 3.0), Micropro's WORDSTAR (3.0) or WORD PERFECT (4.0), the QUALPRO output may serve either as the foundation for writing a report or as supplemental data for a report being written.

In summary, QUALPRO is a highly specialized, flexible and automated PRODUCTIVITY TOOL for Qualitative Research. QUALPRO is menu-driven and easy to learn and use. The software comes with a 102 page MANUAL that includes a GUIDED TOUR designed to lead first time users step-by-step through the programs operations.